

Secure Communication through Wireless-Powered Friendly Jamming: Jointly Online Optimization over Geography, Energy and Time

Pan Zhou, *Member, IEEE*

Abstract—Exploring the interference-emitting friendly jammers to protect the sensitive communications in the presence of eavesdroppers has increasingly being investigated in literature. In parallel, scavenging energy from ambient radio signals for energy-constrained devices, namely *wireless energy harvesting* (WEH), has also drawn significant attention. Without relying on external energy supply, the wireless-powered friendly jammer by WEH from legitimate wireless devices is an effective approach to prolong their lifetime and gain the flexibility in deployments. This paper studies the online optimization of the placement and WEH of a set of friendly jammers in a geographic location with the energy-efficiency (EE) consideration. We adopt a simple “time switching” protocol where power transfer and jammer-assisted secure communications occur in different time blocks when WEH requests are launched. Our scheme has the following important advantages: 1) The proposed online jammers placement and interfering power allocation to attack eavesdroppers is the first distributed and scalable solutions within any specified geographic region; 2) We model the WEH for jammers as a *JAM-NET lifetime maximization problem*, where online scheduling algorithms with heterogeneous energy demands of each jammer (from energy sources) are designed; 3) Under our model, the problem of placing a minimum number of jammers with distance-based power assignments is NP-hard, and near optimal PTAS approximation algorithms are provided; 4) When durations of the eavesdropping and legitimate communicating are available and the scenario is extended to the multi-channels setting, our results are strengthened to see further improved EE and reduced number of jammers. Simulations back up our theory.

Index Terms—Wireless power transfer, energy harvesting, friendly jamming, security, energy efficiency, online learning.

I. INTRODUCTION

A. Background and Motivation

Due to the inapplicability and high computational complexity of cryptography in many dynamic wireless environments, physical layer security techniques [1], [2] for securing the transfer of highly sensitive information in wireless communications have attracted significant attention in the past decades. Systems such as mobile personal healthcare records [5], contactless payment cards [6], telemedicine systems using wireless networks [7] and military sensor networks [8] all employ wireless technologies to transmit potentially sensitive information. In particular, placing jammers as cooperative communication nodes has recently been explored as an effective means to achieve the secure wireless communications

from eavesdroppers [3], [4]. By exploiting the shared nature of wireless channels, the successful deployments of jammers for security must achieve the twin goals that i) reducing the Signal-to-Interference-plus-Noise Ratio (SINR) of eavesdroppers to a level that far below a threshold for successful reception, and ii) maintaining the sufficient channel qualities such that the SINR at the legitimate receivers are not reduced too much so as to prevent the reception of *wireless information transfer* (WIT). However, this is often realized at the expense of additional power consumption for friendly jammers.

Conventional energy harvesting methods rely on various renewable energy sources in the environments, such as solar, wind, vibration and thermoelectric, that are usually unstable and uncontrollable. In contrast, the recent advance in radio frequency (RF) enabled *wireless power transfer* (WPT) technology provides an attractive solution by powering wireless nodes with continuous and stable energy over the air [15]. The key idea of this technology is by leveraging the far-field radiative properties of electromagnetic wave (EMW), the wireless nodes could capture EMW remotely from RF signals and convert it into direct current to charge its battery.

Recently, the WPT has attracted great interests in the research community on energy constrained wireless networks. In [16]–[19], the authors studied the sources simultaneously performing the WPT and WIT to destinations and problem that how the wireless nodes makes use of the harvested energy from WPT to enable communications. Motivated by these works, the process of WET can be fully controlled, hence it is preferred to be applied in wireless networks with critical quality-of-service requirements, such as secure wireless communications. In [20]–[22], the authors considered secure communications with the existence of a single information receiver and several wireless energy-harvesting eavesdroppers. In [23], the authors presented the coexistence of three types of destination in a simple wireless communication scenario: an information receiver, an eavesdropper and a harvesting wireless energy receiver. As noticed, all these works [20]–[23] on only focus on the process of energy-harvesting, the use of which at the receivers (e.g., friendly jammers) for secure communications is not studied. Recently, work [24] used the harvested power at a friendly jammer as a useful resource to emit constructive interference to attack the eavesdroppers that secures the legitimate wireless communication link for the first time. However, their focus is only on a single communication link, where the placement of multiple friendly jammers in a geographical locations and and the energy efficiency (EE)

Pan Zhou is from the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China. E-mail: zhoupannewton@gmail.com

issue in the power management for general wireless networks applications are not studied yet.

Similar to [9]–[11], we consider the following typical scenario that the legitimate communication is often conducted in some restricted geographic locations, where jammers placed in the vicinity are used to secure the legitimate communication. Different from offline (centralized) solution [9]–[11] that the transmission power and number of jammers are required to be optimized separately over a known geographic region, we do not restrict the scale and geometry of the geographic locations and emphasize the potential dynamics of nodes within the secure communication. Hence, distributed and online jammers placement protocols are desirable. In this case, the friendly jammers have the flexibility to be placed randomly at any feasible locations. Therefore, their lifetime is usually constrained to the energy stored in the battery, and WEH as a promising approach [12]–[14] is demanding to prolong their lifetime. Thus, it is highly motivating to study the EE by leveraging both the WEH and the interfering power allocation processes in the power management.

B. Our Work and Contribution

As illustrated in Fig. 1, the legitimate communications within a region, named as *storage* \mathcal{S} , surrounded by a *fence* \mathcal{F} , and the jammers are placed between the space of \mathcal{S} and \mathcal{F} to protect the legitimate communications from eavesdroppers lying outside the fence. To this end, the friendly jammers act as passive security assistants of legitimate communication links. The deployments of this low cost and simple passive jammers brings both the important advantages and challenges: on the one hand, the WEH-based scheme without any power line connection facilitate the flexibly online jammer deployments. Such placements are inherently local and particularly useful in distributed deployments, which is highly desirable for complex geographic areas, e.g., lofty and rugged hills, rough grounds, pot-holed city streets and architectures, etc. and large-scale network deployments; on the other hand, jammers should have low design cost and complexity as well as have high efficiency in energy harvesting method to enable its functionality. Moreover, the jammers are not capable to communicate with each other and can only passively report their “remaining energy status” (as “energy demands” from the perspective of legitimate networks) periodically to the transmitters. In these settings, when a request of placing a new jammer targeting on a passive eavesdropper (or potential eavesdropping position) arrives, an online algorithm needs to decide whether to accept the request and assign a transmission power (one out of F channels in the multi-channel setting) to it. Decisions about the acceptance as well as the power and channel assignments cannot be revoked later.

To solve the above secure communications problem, we propose to use a set of wireless-powered friendly jammers as a defensive and constructive interference-emitting companions, where jammers harvest energy via WPT from the legitimate source nodes. The energy harvesting circuit of the jammers (e.g., consisting of a passive low-pass filter and diode(s) [26]) is very simple and cost effective, and such a configuration is

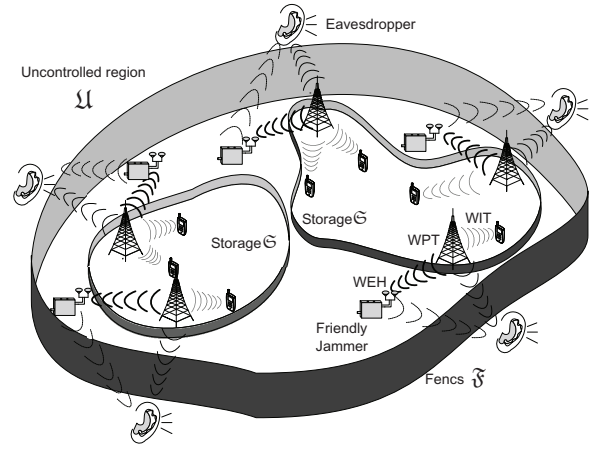


Fig. 1: Friendly jammers-assisted secure communications in the geographic region against eavesdroppers outside the fence. Jammers are placed between the fence \mathcal{S} and the storage \mathcal{F} .

very easy to be controlled by the external energy sources. We use a simple “time switching” scheme [19], [27] such that there are two phases within a complete secure communication circle: namely WPT and WIT for secure communication. In the first phase, due to locations varies over time and CSI is not available from the passive jammers to the energy sources and different energy demand of each jammer, it is challenging to find the optimal energy scheduling algorithm for WPT.

In the second phase, using competitive analysis, we study algorithms using distance-based power assignments from each placed jammer to the eavesdropper location(s). Accepted placement request must satisfies constraints on the SINR for both eavesdroppers and legitimate receivers. The aim is to minimize the number of accepted placement requests of jammers. We first focus on the case of a single channel for the request sets with spatial lengths in $[1, \Delta]$ and average minimal duration between legitimate WIT and eavesdropping in $[1, \Lambda]$.

The main contributions of this work are summarized below:

- 1). The novelty of the work lies in the design of the first distributed protocol that provides secure communication in any geographically restricted communication networks using energy-constrained friendly jammers wirelessly powered by legitimate transmitters as energy sources.
- 2). We consider the energy-efficiency (EE) during the whole design. We adopt a time-division based protocol with during the WPT phase, where legitimate transmitters only provide quota total energy \bar{E} within a total WEH micro-slots of T . The problem is named as the *JAM-NET lifetime maximization*. It is formulated by an adaptive constrained integer linear program (ILP) to meet the goal of EE with heterogenous energy demands. We analysis the originally hard problem from several aspects with practical implementation considerations by advanced online learning algorithms, which indicates that the optimal online scheduling algorithms are available. In addition, for constructive jamming power assignment, we studied the linear power assignment based on the previous *distance-based* power assignment policy, which has the advantage of being energy-minimal.
- 3). The friendly jammers only need to know minimal infor-

mation about the communication taking place. They are proactive rather than reactive, requiring no overhead for the legitimate communication nodes and no synchronization amongst themselves. Our protocol supports dynamic behaviors, e.g., mobility, eavesdropping (communicating) completion or addition/removal of nodes, as along the secure communication are restricted to the storage. However, our proposed protocol is adaptive to the situations such information is available, e.g., exact positions and frequency of both legitimate communications and eavesdropping behaviors, and foreseen further EE improvements and reduced number of jammers.

4). We indicate that it is NP-hard to minimize the number of placed jammers necessary to protect the geographic domain of secure communications.

- We derive, for any fixed ε , the upper bound of $O((1 + \varepsilon)\Delta^{d/2})$ and $O((1 + \varepsilon)\Lambda\Delta^{d/2})$ on the competitive ratio of any deterministic online algorithm without and with the knowledge of duration Λ .
- Then, we extend the result to the general polynomial power assignment with parameter r that cannot yield a competitive ratio worsen than $O((1 + \varepsilon)\Delta^{\min\{r, 1-r\}})$; for the square root power assignment, it yields an upper bound of $O((1 + \varepsilon)\Delta^{d/2})$. In fact, we show that this bound holds for any distance-based power assignment.
- Our upper bounds reveals an exponential gap of the achievable approximation guarantees between deterministic online and offline [9]–[11] algorithms. The main difficulty of the online scenario turns out to be that the request cannot be ordered by length due to distributed deployments. Given $r \in [0, 1]$, we showed that the square root $r = 1/2$ achieves near optimal competitive ratio among all distance-based power assignments and it superior to any other polynomial power assignment.
- We extended our analysis to the multi-channel cases. We generalize the analysis of MULTI-CHAN JAM-Distance algorithm from 1 to F channels. Using $F = F' \cdot F''$ channels is only $\Omega((1 + \varepsilon)F \cdot \Lambda^{1/F'} \cdot \Delta^{1/F''})$ -competitive. It indicates an exponential reduction in the competitive ratio, which indicates that the *multi-channel diversity* could improve the security of legitimate communications.

C. Related Work

Most related work, e.g. [16]–[23], focused on the wiretap channel [28] in the field of information theory, in which a single eavesdropper tries to listen to legitimate communication between a pair of nodes. It is shown that perfect security is possible when the eavesdropper's channel quality is lower than a threshold. Recent works [29], [30] also focus on the MIMO wiretap channel where the transmitter, receiver and eavesdropper may configured with multiple antennas. In [27], the authors had used a wireless-powered relay to help the point-to-point communication. In [31], the authors studied the friendly jamming signal design to help the secure communication based on the knowledge of the uncontrollable energy harvesting process. Different from [27] [31], authors in [24] considered the WEH at a friendly jammer to emitting constructive jamming power for a secure communication link,

where the jamming power and rate parameters are optimized for secure communication. However, most of these works primarily targeted to the theoretical significant due to the simple scenario under consideration but do not explore the geometry of the problem sufficiently.

Vilela et al. [32] showed that without any assumptions on the locations of friendly jammers and eavesdroppers, jammers could co-transmitting with the legitimate transmitter and in the vicinity of a common destination. The authors formulated this setting as a graph and use ILP to find an optimal subset of jamming nodes. In [33], the authors study the asymptotic behaviors for jammers and eavesdroppers at the stochastically distributed locations. In particular, they proposed the concept of *Secure Throughput*, which is based on the probability that a message is successfully received only by legitimate receivers. To our best knowledge, [9]–[11] are the only works that adapt friendly jammers into complex geometric positioning constraints. They provided offline optimal solutions to jammer placement problem that involves both continuous aspects [10], [11] (i.e., power allocations) and discrete aspects [9] (i.e., jammers placements), but they are necessarily to be solved separately. Moreover, all the above works do not consider the issues of WEH and EE for friendly jammers to prolong their lifetime, which are our main focuses. Another important line of this work is the study of competitive ratio of the admitted friendly jammers with instant power allocation in the distributed setting for secure wireless communications for the first time. We note that existing related works on distributed scenarios only studied the competitive ratio for capacity maximization [35] and online admission control [36] in classic wireless communications.

The rest of this paper is organized as follows: Section II describes our system model. Section III proposes the JAM-NET lifetime maximization problem, we analyze its learning performance in several typical and practical implementations. Section IV focuses on the distributed online jammer placement and power allocation problem with competitive analysis. We extend our results to the requests with duration and multi-channel scenarios in Section V. Simulation results are presented in Section VI. The paper is concluded in Section VII.

II. SYSTEM MODELS

A. Environment Model

We consider a Storage/Fence environment model in which legitimate communication takes place within an enclosure specified by one or more polygonal regions $\mathfrak{S} \subset \mathbb{R}^3$, called the storage. We do not assume any knowledge of the locations of communication links in \mathfrak{S} , but we do assume some properties of legitimate communication described below. At first, the legitimate transmitters and receivers can be located at any point $p_s \in \mathfrak{S}$. Further, there exists a *controlled region*, $\mathfrak{C} \subset \mathbb{R}^3$, like the band region that contains \mathfrak{S} in Fig. 1, where there is no eavesdropper able to be within the interior of \mathfrak{C} . The boundary $\partial\mathfrak{C}$ is referred to as the fence \mathfrak{F} . Outside of \mathfrak{C} is the *uncontrolled region* \mathfrak{U} . We assume that there is no hole in \mathfrak{C} , which is a union of simply connected regions; otherwise it can be divided into different regions of storages and fences. We

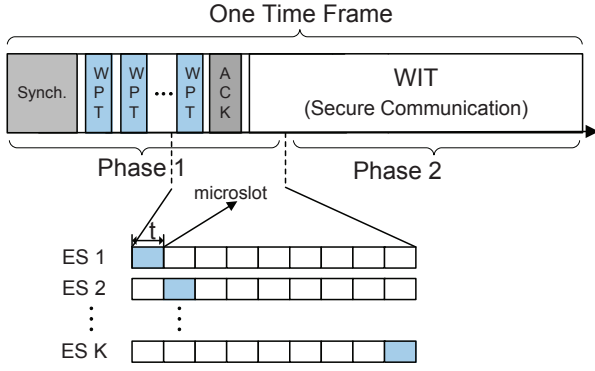


Fig. 2: Time frame divided into two phases. The phase one is for scheduling WPT among multiple energy sources (ESs) and the phase two is for WIT for secure communications.

are necessary to place the jammers distributively in a region \mathcal{A} called the *allowable region*. The allowable region permits us to place the jammers at potential restrictions on locations that belong to the region of \mathcal{C}/\mathcal{S} , e.g., the guarded distance for secure communications or locations that are easily reached for maintenance purpose.

B. Wireless Power Transfer and Energy Harvesting Model

We consider the WPT problem with a set of energy sources (ESs) $\mathcal{X} = \{1, 2, \dots, k, \dots, K\}$ as legitimate transmitters (TxS) from the legitimate communication to power a set of friendly jammers $\mathcal{A} = \{1, 2, \dots, j, \dots, J\}$ with maximal cardinality J under all observed time frames. The ESs are assumed to have no restriction to supply the energy request to each jammer, but the served energy must be used energy-efficiently. Since ESs are geographically randomly distributed and the stochastic placements of jammers, it is desirable to design an online WPT scheduling algorithm (scheduler) from ESs to jammers for WEH over time. We employ the time-switching protocol as illustrated in Fig. 2, where a communication frame is divided into the two phases. The phase of WPT is divided into three sub-phases: i) perform a time synchronization among all ESs to avoid con-channel interference; ii) using the idea of time-division multiplexing, the time block of second phase is further divided into micro slots. Online algorithms are necessary, which schedule the WEH from a single ES k a single jammer j at each microslot (short as 'slot') t based on the observed CSI and energy demands from jammers to ESs (detailed below); and iii) send an ACK to indicate the completion of the WPT phase. The jammers can also initiate the next frame for WEH at the end of last frame when their remaining energy-level is lower than some threshold. Hence the WIT phase is not necessarily to be launched at each time frame.

Let the equivalent complex baseband channel from a ES k to a jammer j is denoted by $g_{k,j}^t(v)$, where v denotes the fading state of the CSI with the instant channel power gain $h_{k,j}^t(v) = |g_{k,j}^t(v)|^2$ at a micro slot t . For each fading state v , the portion of signal power split to secure communication is denoted by $\alpha(v)$ with $0 \leq \alpha(v) \leq 1$, and that to WEH as $1 - \alpha(v)$. In general the $\alpha(v)$ can be adjusted over different fading state over time. For the WPT phase, the scheduling strategies contains the following two cases, **Case I**: $h_{k,j}^t(v)$

is perfectly known at ESs for each fading state v , referred to as the *known CSI at ESs*, which is a simple model but not realistic; **Case II**: $h_{k,j}^t(v)$ is unknown at the ESs (TxS) for all the fading state v , referred to as *unknown CSI at ESs*, which is practical problem as the passive jammers usually are not capable to communicate with ESs to estimate the CSI. In this case, we need to estimate the mean values of $h_{k,j}^t(v)$ by online learning algorithms; Given $\alpha(v)$, the harvested energy (HE) at each slot t (normalized to unitary) at each fading state v can be expressed as

$$Q_{k,j}(v) = \xi(1 - \alpha(v))h_{k,j}^t(v)p_{k,j}^t(v), \quad (1)$$

where ξ is a constant coefficient that accounts for the loss in the energy transducer for converting the HE to electrical energy to be stored, and $p_{k,j}^t(v)$ denotes the scheduled energy from ES k to j at slot t . Denoted each jammer has an energy demand $c_{X_t,j}$ at slot t , we have the scheduled energy request from ES X_t at t such that

$$p_{k,j}^t(v) = c_{X_t,j}, X_t = k. \quad (2)$$

Then, the expected HE at each jammer is then given by

$$\mu_{k,j} = \mathbb{E}_v[Q_{k,j}(v)], \forall k \in \mathcal{X}, j \in \mathcal{J}. \quad (3)$$

Based on the time-division scheme, at each slot t for WEH, only one ES is activated for WPT according to its distance to the boundary of storage \mathcal{S} , based on which the appearance of an ES X_t is normalized as an independently with identical distributed $\mathbb{P}\{X_t = k\} = \pi_k, k \in \mathcal{X}$. In this scheme, each scheduled jammer $j \in \mathcal{A}$ generates a non-negative HE (as reward) $Y_{j,t}$. W.l.o.g., under a given ES $X_t = k$, the HE $Y_{j,t}$'s are independent random variables in $[0, 1]$, where the better channel gain $h_{k,j}^t(v)$ (e.g., schedule the nearer jammer) and the larger energy demand provide better HE at jammer j . But, the conditional expectation $\mathbb{E}[Y_{j,t}|X_t = k] = u_{k,j}$ is unknown to the scheduler. Moreover, an energy demand is realized if jammer j is scheduled under ES k . We consider fixed and known energy demands in this paper, where the $c_{k,j} > 0$ when jammer j is served by ES k .

We can formulate the problem as a constrained context-bandit (CMAB) [34] online learning problem. Similar to traditional contextual bandits, the ES X_t as a *context* is observable at the beginning t , while only the HE of the jammer taken by the scheduler is revealed at the end of slot t . Specifically, at the beginning of slot t , the scheduler observes the context X_t and takes a jammer A_t from $\{0\} \cup \mathcal{A}$, where "0" represents a *dummy* jammer that the scheduler skips the current context. Let Y_t and Z_t be the HE and energy demand received for the scheduler in slot t , respectively. If the scheduler takes a jammer $A_t = j > 0$, then the HE is $Y_t = Y_{j,t}$ and the energy demand is $Z_t = c_{X_t,j}$. Otherwise, if the scheduler takes the dummy jammer $A_t = 0$, neither HE nor energy demand is incurred, i.e., $Y_t = 0$ and $Z_t = 0$. We focus on the CMAB in this work with a known time-horizon T and limited energy budget \bar{E} for the goal of EE in WPT, where the process ends when the scheduler runs out of the energy budget or at the end of time T .

Formally, an online learning algorithm F is a function that maps the historical observations $\mathcal{H}_{t-1} = (X_1, A_1, Y_1; X_2, A_2, Y_2; \dots; X_{t-1}, A_{t-1}, Y_{t-1})$ and the current ES X_t to a jammer $A_t \in \{0\} \cup \mathcal{A}$. The objective of the

algorithm is to maximize the expected total HEs $U_F(T, \bar{E})$ for a given T and an energy budget \bar{E} , i.e.,

$$\begin{aligned} & \text{maximize } U_F(T, \bar{E}) = \mathbb{E}_F \left[\sum_{t=1}^T Y_t \right] \\ & \text{subject to } \sum_{t=1}^T Z_t \leq \bar{E}, \end{aligned}$$

where the expectation is taken over the distributions of ESs and energy rewards. Note that we consider a “hard” energy budget constraint, i.e., the total energy demands should not be greater than \bar{E} under any realization.

We measure the performance of the algorithm F by comparing it with the optimal one, which is the optimal algorithm with known statistics, including the knowledge of π_k 's, and $u_{k,j}$'s and $c_{k,j}$'s. Let $U^*(T, \bar{E})$ be the expected total HE for the jammers' networks (JAM-NET) obtained by an offline optimal algorithm (with hindsight full knowledge). Then, we define the term “*regret*” of the algorithm Γ as

$$R_\Gamma(T, \bar{E}) = U^*(T, \bar{E}) - U_\Gamma(T, \bar{E}).$$

The objective of F is then to minimize the regret. We are interested in the asymptotic performance, where the time-horizon T and the energy budget \bar{E} grow to infinity proportionally, i.e., with a fixed ratio $\rho = \bar{E}/T$. The coordination process of ESs to use up the total budget \bar{E} is discussed in Section III.

C. Secure Communication Model

The communication model is similar to that of [9]–[11]. W.l.o.g., we assume that the transmission power \tilde{P} is the same for all legitimate transmitters and the receiving power \bar{P} is nearly the same for all receivers within \mathfrak{S} . On the other hand, the heard signal at eavesdropper suffers path loss with the path-loss exponent γ . Formally, for an eavesdropper p_e listens to a transmitter $p_s \in \mathfrak{S}$; the received power is $\bar{P} d_{ps}^{-\gamma}$, where γ typically in range $[2, 6]$ and $d_{pq} = \|p - q\|$ is the Euclidean distance between p and q .

We assume that it is co-channel interference free among legitimate communications and we use the *Signal-to-interference Ratio* (SIR) as physical model that only jamming signals cause interference. Formally, for a legitimate receiver p_s in \mathfrak{S} , it is the ratio of the transmitted signal power to the total interference contributed by the jammers,

$$\text{SIR}(J, p_s) = \frac{\tilde{P}}{\sum_{j \in J} P_j d_{jp_s}^{-\gamma}}, \quad (4)$$

where P_j is the transmission power of jammer j . Similarly, for an eavesdropper p_e , the transmission signal from the storage suffers path loss. In the case the jammer is passive, we make an observation point that the maximal signal power received from a transmitter at the nearest location of p_e on \mathfrak{S} is $s(p_e)$, and we define

$$\text{SIR}(J, p_e) = \frac{\tilde{P} d_{s(p_e)p_e}^{-\gamma}}{\sum_{j \in J} P_j d_{jp_e}^{-\gamma}}. \quad (5)$$

In [9]–[11], the authors use a simple truth that the total interference at a location p is usually dominated by the interference from the nearest jammer to p due to the received power decreasing exponentially with distance. However, our model gives religious interference analysis that contributed by all transmitters.

To facilitate successful receptions of legitimate transmissions, we require that $\text{SIR}(J, p_s) > \delta_s$, for all points $p_s \in \mathfrak{S}$ and some specified value of δ_s . Similarly, to make the eavesdroppers unable to receive secure messages, we require that $\text{SIR}(J, p_e) < \delta_e$, for points $p_e \in \mathbb{R}^2/\mathfrak{C}$ and some parameter δ_e . In summary, the set of friendly jammers need to satisfy the following constraints:

$$\begin{aligned} & \text{SIR}(J, p_s) > \delta_s, \forall p_s \in \mathfrak{S}, \\ & \text{SIR}(J, p_e) < \delta_e, \forall p_e \in \mathbb{R}^2/\mathfrak{C}, \end{aligned} \quad (6)$$

Finally, as indicated in [10], jamming the eavesdroppers at the fence \mathfrak{F} is sufficient to ensure that eavesdroppers located outside the fence are also jammed successfully; ensuring the legitimate receiver on the boundary of \mathfrak{F} is not jammed is sufficient to guarantee that receivers inside \mathfrak{F} are not jammed.

III. JAM-NET LIFETIME MAXIMIZATION: CONSTRAINED CONTEXTUAL BANDITS WITH HETEROGENEOUS COSTS

We name the energy harvest problem of JAM-NET as the *JAM-NET lifetime maximization* problem. We discuss the design of algorithms in heterogeneous-energy-demand systems where energy demand $c_{k,j}$ depends on k and j . We summary the main results here first and then go into details.

To schedule the WPT from multiple ESs to jammers, we show that significant complexity incurs when making decisions under those “medium-quality” ESs other the best and worst ESs. Because the scheduler needs to balance between the instantaneous expected harvest energy rewards and the future rewards, which is very difficult due the coupling effect introduced by the time and energy budgets constraints. Thus, we resort to approximations by relaxing the hard time and energy budget constraints to an average energy budget constraint. Now, the problem becomes to maximize the expected rewards (of total harvest energy) with average budget constraint \bar{E}/T . However, the problem of using the fixed average budget \bar{E}/T and not taking the remaining time τ and remaining energy budget b_τ into account would not explore the dynamic structure of the problem, which will lead to suboptimal solutions. Hence, we propose an Adaptive Integer Linear Program (AILP) that replaces the fixed average budget \bar{E}/T by the *remaining average budget*, i.e., b_τ/τ . We indicate that the performance analysis is non-trivial for the AILP, although the intuition behind the formulation is quite natural. Using concentration inequalities, we can show that b_τ/τ under AILP concentrates near the average budget \bar{E}/T with high probability. This proves that the proposed AILP algorithms achieves an expected total HE only within a constant to the optimum expect for certain boundary cases.

Next, gaining the insight from the case of the known statistics of the expected HE for each jammer, we extend our analysis to that the expected HE as rewards for jammers are unknown. During the analysis, we find that the AILP algorithm only require the ranking of the expected rewards rather than their actual values. Inspired by this, we combine the famous upper-confidence-bound (UCB) algorithm [38] and show that, for the general heterogeneous-energy-demand systems, our proposed ϵ -ESs-JamNet-UCB-AILP algorithm achieves $O(\log(t))$ regret that is learning-rate optimal except for certain boundary case, where it achieves $O(\sqrt{t})$ regret.

A. Approximation of the Offline Optimal Algorithm: Known CSI at ESs

In this subsection, we first study the case with known statistics. We consider the case where the energy demand for each jammer k under context j is fixed at $c_{k,j}$, which may be different for different k and j .

With known statistics, the scheduler knows the context (ESs) distribution π_k 's, the energy demands $c_{k,j}$'s, and the expected rewards of HEs $u_{k,j}$'s. With heterogeneous energy demands, the quality of the WEH process a jammer j under a context k is roughly captured by its *normalized expected HE*, defined as $\eta_{k,j} = u_{k,j}/c_{k,j}$. However, the scheduler cannot only focus on the “best” jammer to serve, i.e., $j_k^* = \arg\max_{j \in \mathcal{A}} \eta_{k,j}$, for context ES k . This is because there may exist another jammer j' such that $\eta_{k,j'} < \eta_{k,j_k^*}$, but $u_{k,j'} > u_{k,j_k^*}$ (and surely, $c_{k,j'} > c_{k,j_k^*}$). If the energy budget allocated to each ES k out of the total is sufficient, then the scheduler may take jammer j' to maximize the expected reward of HE. Therefore, the ALP algorithm in this case needs to decide the probability to take jammer j under ES k , by solving an ILP problem with an additional constraint that only one jammer can be taken under each ES. We can show that ALP achieves $O(1)$ regret in non-boundary cases, and $O(\sqrt{T})$ regret in boundary cases. We note that the regret analysis of ALP in this case is much more difficult due to the additional constraint that couples all jammers under each ES.

In this case, the scheduler needs to consider all jammers under each ES. Let $p_{k,j}$ be the probability that jammer j is taken under ES k . We define the following ILP problem:

$$(\mathcal{LP}'_{T,\bar{E}}) \quad \text{maximize} \quad \sum_{k=1}^K \pi_k \sum_{j=1}^J p_{k,j} u_{k,j}, \quad (7)$$

subject to:

$$\sum_{k=1}^K \pi_k \sum_{j=1}^J p_{k,j} c_{k,j} \leq \bar{E}/T, \quad (7a)$$

$$\sum_{j=1}^J p_{k,j} \leq 1, \forall k, \quad (7b)$$

$$p_{k,j} \in [0, 1]. \quad (7c)$$

The above LP problem $\mathcal{LP}'_{T,\bar{E}}$ can be solved efficiently by optimization tools. Let $\hat{v}(p)$ be the maximum value of $\mathcal{LP}'_{T,\bar{E}}$. Similar to Lemma 1, we can show that $T\hat{v}(p)$ is an upper bound of the expected total HE, i.e., $T\hat{v}(p) \leq U^*(T, \bar{E})$.

To obtain insight from the solution of $\mathcal{LP}'_{T,\bar{E}}$, we derive an explicit representation for the solution by analyzing the structure of $\mathcal{LP}'_{T,\bar{E}}$. Note that there are two types of (non-trivial) constraints in $\mathcal{LP}'_{T,\bar{E}}$, one is the “inter-ES” energy budget constraint (7a), the other is the “intra-ES” constraint (7b). These constraints can be decoupled by first allocating energy budget for each context, and then solving a subproblem with the allocated energy budget constraint for each ES. Specifically, let ρ_k be the energy budget allocated to ES k , then $\mathcal{LP}'_{T,\bar{E}}$ can be decomposed as follows:

$$\begin{aligned} & \text{maximize} \quad \sum_{k=1}^K \pi_k \hat{v}_k(\rho_k), \\ & \text{subject to:} \quad \sum_{k=1}^K \pi_k \rho_k \leq \bar{E}/T \end{aligned}$$

where

$$\begin{cases} (\mathcal{SP}_k) \quad v_k(\rho_k) = \text{maximize} \quad \sum_{j=1}^J p_{k,j}, & (8a) \\ \text{subject to} \quad \sum_{j=1}^J p_{k,j} c_{k,j} \leq \rho_k, & (8b) \\ \sum_{j=1}^J p_{k,j} \leq 1, p_{k,j} \in [0, 1]. & (8c) \end{cases}$$

Next, by analyzing sub-problem \mathcal{SP}_k , we show that some jammers can be deleted without affecting the performance, i.e., the probability is 0 in the optimal solution.

Lemma 1. For any given $\rho_k \geq 0$, there exists an optimal solution of \mathcal{SP}_k , i.e., $\mathbf{p}_k^* = (p_{k,1}^*, p_{k,2}^*, \dots, p_{k,J}^*)$, satisfies:

- (1) For j_1 , if there exists another jammer j_2 , such that $\eta_{k,j_1} \leq \eta_{k,j_2}$ and $u_{k,j_1} \leq u_{k,j_2}$, then $p_{k,j_1}^* = 0$;
- (2) For k_1 , if there exists two jammers j_2 and j_3 , such that $\eta_{k,j_2} \leq \eta_{k,j_1} \leq \eta_{k,j_3}$, $u_{k,j_2} \leq u_{k,j_1} \leq u_{k,j_3}$, and $\frac{u_{k,j_1} - u_{k,j_3}}{c_{k,j_1} - c_{k,j_3}} \leq \frac{u_{k,j_2} - u_{k,j_3}}{c_{k,j_2} - c_{k,j_3}}$, then $p_{k,j_1}^* = 0$.

Intuitively, the first part of Lemma 1 shows that if a jammer has small normalized and original expected HE, then it can be removed. The second part of Lemma 1 shows that if a jammer has small normalized expected HE and medium original expected HE, but the increasing rate is smaller than another jammer with larger expected HE, then it can also be removed.

Proof: The key idea of this proof is that, if the conditions is satisfied, and there is a feasible solution $\mathbf{p}_k = (p_{k,1}, p_{k,2}, \dots, p_{k,J})$ such that $p_{k,j_1} > 0$, then we can construct another feasible solution \mathbf{p}'_k such that $p'_{k,j_1} = 0$, without reducing the objective value $v_k(\rho_k)$.

We first prove part (1). Under the conditions of part (1), if \mathbf{p}_k is a feasible solution of \mathcal{SP}_k with $p_{k,j_1} > 0$, then consider another solution \mathbf{p}'_k , where $p'_{k,j} = p_{k,j}$ for $j \notin \{j_1, j_2\}$, $p'_{k,j_1} = 0$, and $p'_{k,j_2} = p_{k,j_2} + p_{k,j_1} \min\{\frac{c_{k,j_1}}{c_{k,j_2}}, 1\}$. Then, we can verify that \mathbf{p}'_k is a feasible solution of (\mathcal{SP}_k) , and the objective value under \mathbf{p}'_k is no less than that under \mathbf{p}_k .

For the second part, if the conditions are satisfied and $p_{k,j_1} > 0$, then we construct a new solution \mathbf{p}'_k by re-allocating the energy budget consumed by jammer j_1 to jammer j_2 and j_3 , without violating the constraints. Specifically, we set the probability the same as the original solution for other jammers, i.e., $p'_{k,j} = p_{k,j}$ for $j \notin \{j_1, j_2, j_3\}$, and set $p'_{k,j_1} = 0$ for jammer j_1 . For j_2 and j_3 , to maximize the objective function, we would like to allocate as much energy budget as possible to j_3 unless there is remaining energy budget. Therefore, we set $p'_{k,j_2} = p_{k,j_2}$ and $p'_{k,j_3} = p_{k,j_3} + \frac{p_{k,j_1} c_{k,j_1}}{c_{k,j_3}}$, if $\sum_{j \neq j_1} p_{k,j} + \frac{p_{k,j_1} c_{k,j_1}}{c_{k,j_3}} \leq 1$; or, $p'_{k,j_2} = p_{k,j_2} + \frac{p_{k,j_1} c_{k,j_1} - (1 - \sum_{j \neq j_1} p_{k,j}) c_{k,j_3}}{c_{k,j_2} - c_{k,j_3}}$ and $p'_{k,j_3} = p_{k,j_3} + \frac{(1 - \sum_{j \neq j_1} p_{k,j}) c_{k,j_2} - p_{k,j_1} c_{k,j_1}}{c_{k,j_2} - c_{k,j_3}}$, if $\sum_{j \neq j_1} p_{k,j} + \frac{p_{k,j_1} c_{k,j_1}}{c_{k,j_3}} > 1$.

We can verify that \mathbf{p}_k satisfies the constraints of (\mathcal{SP}_k) but the objective value is no less than that under \mathbf{p}_k . ■

With Lemma 1, the scheduler can ignore some jammers that will obviously be allocated with zero probability under a given context k . We call the set of the remaining jammers as

candidate set for context k , denoted as \mathcal{A}_k . We propose an algorithm to construct the candidate jammer set for context k , as shown in Algorithm 1.

Algorithm 1 Find Candidate Jammer Set for ES k

Input: $c_{k,j}$'s, $u_{k,j}$'s, for all $1 \leq j \leq J$;

Output: \mathcal{A}_k ;

- 1: Calculate normalized HE as rewards: $\eta_{k,j} = u_{k,j}/c_{k,j}$;
- 2: Sort jammers in descending order of their normalized rewards of HEs:

$$\eta_{k,j_1} \geq \eta_{k,j_2} \geq \dots \geq \eta_{k,j_J}.$$

- 3: **for** $a = 2$ **to** J **do**
 - 4: **if** $\exists a' < a$ such that $u_{k,j_a} \leq u_{k,j_{a'}}$ **then**
 - 5: $\mathcal{A}_k = \mathcal{A}_k / \{j_a\}$;
 - 6: **end if**
 - 7: **end for**
 - 8: $a = 1$;
 - 9: **while** $a \leq J - 1$ **do**
 - 10: Find the jammer with highest increasing rate:

$$a^* = \operatorname{argmax}_{a': a' > a, j_{a'} \in \mathcal{A}_k} \frac{u_{k,j_{a'}} - u_{k,j_a}}{c_{k,j_{a'}} - c_{k,j_a}}.$$
 - 11: Remove the jammers in between:

$$\mathcal{A}_k = \mathcal{A}_k / \{j_{a'} : a < a' < a^*\}.$$
 - 12: Move to the next candidate jammer: $a = a^*$;
 - 13: **end while**
-

For ES k , assume that the candidate jammer set $\mathcal{A}_k = \{j_{k,1}, j_{k,2}, \dots, j_{k,J_k}\}$ has been sorted in descending order of their normalized energy rewards, i.e., $\eta_{k,j_{k,1}} \leq \eta_{k,j_{k,2}} \leq \dots \leq \eta_{k,j_{k,J_k}}$. From Algorithm 1, we know that $u_{k,j_{k,1}} < u_{k,j_{k,2}} < \dots < u_{k,j_{k,J_k}}$, and $c_{k,j_{k,1}} < c_{k,j_{k,2}} < \dots < c_{k,j_{k,J_k}}$.

The scheduler now only needs to consider the jammers in the candidate set \mathcal{A}_k . To decouple the “intra-ES” constraint (7b), we introduce the following transformation:

$$p_{k,j_k,a} = \begin{cases} \tilde{p}_{k,j_k,a} - \tilde{p}_{k,j_k,a+1}, & \text{if } 1 \leq a \leq J_k - 1, \\ \tilde{p}_{k,j_k,J_k}, & \text{if } a = J_k, \end{cases} \quad (9)$$

where $\tilde{p}_{k,j_k,a} \in [0, 1]$, and $\tilde{p}_{k,j_k,a} \leq \tilde{p}_{k,j_k,a+1}$ for $1 \leq a \leq J_k - 1$. Substituting the transformations into (\mathcal{SP}_k) and reformulate it as

$$(\widetilde{\mathcal{SP}}_k) \text{ maximize } \sum_{a=1}^{J_k} \tilde{p}_{k,j_k,a} \tilde{u}_{k,j_k,a},$$

subject to:

$$\begin{cases} \sum_{a=1}^{J_k} \tilde{p}_{k,j_k,a} \tilde{c}_{k,j_k,a} \leq \rho_k, & (10a) \\ \tilde{p}_{k,j_k,a} \leq \tilde{p}_{k,j_k,a+1}, 1 \leq a \leq J_k - 1, & (10b) \\ \tilde{p}_{k,j_k,a} \in [0, 1], \forall a, & (10c) \end{cases}$$

where

$$\tilde{u}_{k,j_k,a} = \begin{cases} u_{k,j_{k,1}}, & \text{if } a = 1, \\ u_{k,j_{k,a}} - u_{k,j_{k,a-1}}, & \text{if } 2 \leq a \leq J_k, \end{cases} \quad (11)$$

$$\tilde{c}_{k,j_k,a} = \begin{cases} c_{k,j_{k,1}}, & \text{if } a = 1, \\ c_{k,j_{k,a}} - c_{k,j_{k,a-1}}, & \text{if } 2 \leq a \leq J_k, \end{cases} \quad (12)$$

Next, we show that the constraint(10) can indeed be removed. For each $j_{k,a}$, we can view $\tilde{c}_{k,j_k,a}$ and $\tilde{u}_{k,j_k,a}$ as the energy demand and expected HE of a virtual jammer.

Let $\tilde{\eta}_{k,j_k,a} = \tilde{u}_{k,j_k,a} / \tilde{c}_{k,j_k,a}$ be the normalized expected HE of virtual jammer $j_{k,a}$. For $a = 1$, using $\frac{u_{k,j_{k,1}}}{c_{k,j_{k,1}}} \geq \frac{u_{k,j_{k,2}}}{c_{k,j_{k,2}}}$, we can show that $\tilde{\eta}_{k,j_{k,1}} \geq \tilde{\eta}_{k,j_{k,2}}$. For $2 \leq a \leq J_k - 1$ using $\frac{u_{k,j_{k,a}} - u_{k,j_{k,a-1}}}{c_{k,j_{k,a}} - c_{k,j_{k,a-1}}} \geq \frac{u_{k,j_{k,a+1}} - u_{k,j_{k,a}}}{c_{k,j_{k,a+1}} - c_{k,j_{k,a}}}$, we can show that $\tilde{\eta}_{k,j_{k,a}} \geq \tilde{\eta}_{k,j_{k,a+1}}$. In other words, we can verify that $\tilde{\eta}_{k,j_{k,1}} \geq \tilde{\eta}_{k,j_{k,2}} \geq \dots \geq \tilde{\eta}_{k,j_{k,J_k}}$. Thus, without constraint (10), the optimal solution $\tilde{\mathbf{p}}_k = [\tilde{p}_{k,j_1}, \tilde{p}_{k,j_2}, \dots, \tilde{p}_{k,j_{J_k}}]$ automatically satisfied $\tilde{p}_{k,j_1}^* \geq \tilde{p}_{k,j_2}^* \geq \dots \geq \tilde{p}_{k,j_{J_k}}^*$. Hence, we can remove the constraint (10), and thus decouple the probability constraint under a ES.

We can thus rewrite the global ILP problem using the above transformations

$$\begin{aligned} (\widetilde{\mathcal{LP}}'_{T,\bar{E}}) \text{ maximize } & \sum_{k=1}^K \sum_{a=1}^{J_k} \pi_k \tilde{p}_{k,j_k,a} \tilde{u}_{k,j_k,a}, \\ \text{subject to } & \sum_{k=1}^K \sum_{a=1}^{J_k} \pi_k \tilde{p}_{k,j_k,a} \tilde{c}_{k,j_k,a} \leq \frac{\bar{E}}{T}, \\ & \tilde{p}_{k,j_k,a} \in [0, 1], \forall k, 1 \leq a \leq J_k. \end{aligned}$$

The solution of $\widetilde{\mathcal{LP}}'_{T,\bar{E}}$ follows a threshold structure. We sort all ES-(virtual-)jammer pairs (k, j_a) in descending order of their normalized expected HEs. Let $k^{(i)}, j^{(i)}$ be the context index and jammer index of the i -th pair, respectively. Namely, $\tilde{\eta}_{k^{(1)},j^{(1)}} \geq \tilde{\eta}_{k^{(2)},j^{(2)}} \geq \dots \geq \tilde{\eta}_{k^{(M)},j^{(M)}}$, where $M = \sum_{k=1}^K J_k$ is the total number of candidate jammers for all ESs. Define a threshold corresponding to $\rho = \bar{E}/T$,

$$\tilde{i}(\rho) = \max\{i : \sum_{i'=1}^i \pi_{k^{(i')}} \tilde{c}_{k^{(i')},j^{(i')}} \leq \rho\}, \quad (13)$$

where $\rho = \bar{E}/T$ is the average energy budget. We can verify that the following solution is optimal for $\widetilde{\mathcal{LP}}'_{T,\bar{E}}$:

$$\tilde{p}_{k^{(i)},j^{(i)}} = \begin{cases} 1, & \text{if } 1 \leq i \leq \tilde{i}(\rho), \\ \frac{\rho - \sum_{i'=1}^{\tilde{i}(\rho)} \pi_{k^{(i')}} \tilde{c}_{k^{(i')},j^{(i')}}}{\pi_{k^{(\tilde{i}(\rho))},j^{(\tilde{i}(\rho))}} \tilde{c}_{k^{(\tilde{i}(\rho))},j^{(\tilde{i}(\rho))}}}, & \text{if } i = \tilde{i}(\rho) + 1, \\ 0, & \text{if } i > \tilde{i}(\rho) + 1. \end{cases} \quad (14)$$

Then, the optimal solution of $\widetilde{\mathcal{LP}}'_{T,\bar{E}}$ can be calculated using the reverse transformation from $\tilde{p}_{k,j}(\rho)$'s to $p_{k,j}(\rho)$'s.

1) *ALP Algorithm:* Obviously, the ALP algorithm replaces the average constraint \bar{E}/T in $\widetilde{\mathcal{LP}}'_{T,\bar{E}}$ with the average remaining energy budget b_τ/τ , and obtains probability $p_{k,j}(b_\tau/\tau)$. Under context k , the ALP algorithm take jammer j with probability $p_{k,j}(b_\tau/\tau)$. Note that the remaining energy budget b_τ does not follow any classic distribution in heterogeneous-energy-demand systems. However, by using the method of averaged bounded differences [37], we show that there has the concentration property holds.

Lemma 2. For $0 < \delta < 1$, there exists a positive number \mathcal{K} , such that under the ALP algorithm, the remaining energy budget b_τ satisfies

$$\begin{aligned} \mathbb{P}\{b_\tau > (\rho + \delta)\tau\} &\leq e^{-\mathcal{K}\delta^2\tau}, \\ \mathbb{P}\{b_\tau < (\rho - \delta)\tau\} &\leq e^{-\mathcal{K}\delta^2\tau}. \end{aligned}$$

Proof: We prove the lemma using the method of averaged bounded differences [37]. The process is similar to Section 7.1 in [37], except that we consider the remaining energy budget and the successive differences of the remaining energy budget are bounded by c_{\max} .

Specifically, let $\tilde{c}_{t'}$, $1 \leq t' \leq T$ be the energy budget consumed under *ALP*, and let $\tilde{\mathbf{c}}_{t'} = (\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_{t'})$. Then the remaining energy budget at slot t (the remaining time $\tau = T - t + 1$), i.e., b_{T-t+1} is a function of $\tilde{\mathbf{c}}_t$. We note that under *ALP*, the expectation of the ratio between the remain energy budget and the remaining time does not change, i.e., for any $b \leq \sum_{j=1}^t \pi_k c_j^*$ (here $c_j^* = \max_k c_{k,j}$), if $b_\tau = b$, then $\mathbb{E}[b_{T-t+1}/(\tau - 1)] = b/\tau$. Thus, we can verify that for any $1 \leq t' \leq t$, we have

$$\mathbb{E}[b_{T-t+1}|\tilde{\mathbf{c}}_{t'}] = b_{T-t'+1} - \frac{b_{T-t'+1}}{T-t'+1}(t-t').$$

Note that $\Delta b = b_{T-t'+2} - b_{T-t'+1} \leq c_{\max}$ and $b_{T-t'+2} \geq -c_{\max}$, we have

$$\begin{aligned} & |\mathbb{E}[b_{T-t+1}|\tilde{\mathbf{c}}_{t'}] - \mathbb{E}[b_{T-t+1}|\tilde{\mathbf{c}}_{t'-1}]| \\ & \leq \max_{0 \leq \Delta b \leq c_{\max}} \left\{ \left| \Delta b - \frac{b_{T-t'+2}}{T-t'+2} \right| \right\} \frac{T-t+1}{T-t'+1} \\ & \leq \frac{2c_{\max}(T-t+1)}{T-t'+1}. \end{aligned}$$

Moreover,

$$\begin{aligned} \sum_{t'=1}^t \left[\frac{2c_{\max}(T-t+1)}{T-t'+1} \right]^2 &= 4c_{\max}^2(T-t+1)^2 \sum_{t'=1}^t \frac{1}{(T-t'+1)^2} \\ &= 4c_{\max}^2(T-t+1)^2 \sum_{\tau'=T-t+1}^T \frac{1}{(\tau')^2} \\ &\approx 4c_{\max}^2(T-t+1)^2 \int_{T-t+1}^T \frac{1}{(\tau')^2} d\tau' \\ &= 4c_{\max}^2(T-t+1) \frac{t-1}{T}. \end{aligned}$$

According to Theorem 5.3 in [37], and noting $\tau = T - t + 1$, $\mathbb{E}[b_\tau] = \rho\tau$, we have

$$\mathbb{P}\{b_\tau > \mathbb{E}[b_\tau] + \delta_\tau\} \leq e^{-\frac{2T(\delta\rho\tau)^2}{4c_{\max}^2(T-t+1)(t-1)}} \leq e^{-\frac{T\delta^2\bar{E}^2\tau}{2c_{\max}^2T^2(t-1)}} \leq e^{-\frac{\delta^2\rho^2}{2c_{\max}^2}}$$

$$\text{and similarly, } \mathbb{P}\{b_\tau < \mathbb{E}[b_\tau] - \delta_\tau\} \leq e^{-\frac{\delta^2\rho^2}{2c_{\max}^2}},$$

Choosing $\mathcal{K} = \frac{\rho^2}{2c_{\max}^2}$ concludes the proof. \blacksquare

Then, using similar methods in Section 3, we can show that the generalized *ALP* algorithm achieves $O(1)$ regret in non-boundary cases, and $O(\sqrt{T})$ regret in boundary cases, where the boundaries are now defined as $Q_i = \sum_{i'=1}^i \pi_{k(i')} \tilde{c}_{k(i'),j(i')}$.

Next, we show that the ϵ -First policy with CLT will achieve $O(\log T)$ regret except for the boundary cases, where it achieves $O(\sqrt{T})$ regret. On one hand, according to Hoeffding-Chernoff bound, if all comparisons pass the confidence level test, then with probability at least $1 - KJ^2T^{-2}$, the algorithm obtains the correct rank and provide a right solution for the problem $(\mathcal{LP}'_{\tau,b})$. On the other hand, because $\Delta^* > 0$, from the analysis in the previous section, we know that the exploration stage will end within $O(\log T)$ rounds with high probability. Therefore, the expected regret is the same as that in the case with known $\Delta_{\min}^{(\epsilon)}$.

Lemma 2 also states that the average remaining budget δ , τ stays in a neighborhood of the initial average budget ρ with high probability. Hence, if the initial average budget ρ is not on boundaries, i.e., the critical values under which the threshold $\tilde{j}(\rho)$ changes, then the probability of threshold changing is bounded. Therefore, we can show that the *ALP* algorithm

achieves a very good performance within a constant distance from the optimum, except for certain boundary cases. Specifically, for $1 \leq k \leq K$, let q_k be the cumulative distribution function, i.e., $q_k = \sum_{k'=1}^k \pi_{k'}$, and w.l.o.g., let $q_0 = 0$. The following theorem states the approximate optimality of *ALP* for the cases where $\rho \neq q_k$ ($k = 1, 2, \dots, K-1$). We note that $k = 0$ and $k = K$ are trivial cases where *ALP* is optimal.

Theorem 3. *Given any fixed $\rho \in (0, 1)$ satisfying $\rho \neq q_k$, $k = 1, 2, \dots, K-1$, the *ALP* algorithm achieves an $O(1)$ regret. Specifically,*

$$U^*(T, \bar{E}) - U_{ALP}(T, \bar{E}) \leq \frac{u_1^* - u_K}{1 - e^{-2\delta^2}},$$

where $\delta = \min\{\rho - q_{\tilde{k}}, q_{\tilde{k}+1} - \rho\}$.

Proof: The proof of this theorem uses the following two facts derived from lemma 2: $\mathbb{E}[v(b_\tau/\tau)] = v(\rho)$ if the threshold $\tilde{j}(b_\tau/\tau) = \tilde{j}(\rho)$ for all possible b_τ 's, and the probability that $\tilde{j}(b_\tau/\tau) \neq \tilde{j}(\rho)$ decays exponentially. Please referred to Appendix B.1 of the supplementary material for details.

When considering the boundary cases, we can show similarly that the *ALP* achieves $O(\sqrt{T})$ regret. \blacksquare

Theorem 4. *Given any fixed $\rho = q_j$, $k = 1, 2, \dots, K-1$, the *ALP* algorithm achieves $O(\sqrt{T})$ regret. Specifically,*

$$U^*(T, \bar{E}) - U_{ALP}(T, \bar{E}) \leq \Theta^{(o)}\sqrt{T} + \frac{u_1^* - u_K}{1 - e^{-2\delta'^2}},$$

where $\Theta^{(o)} = 2(u_1^* - U_K^*)\sqrt{\rho(1-\rho)}$ and $\delta' = \min\{\rho - q_{\tilde{j}(\rho)-1}, q_{\tilde{j}(\rho)+1} - \rho\}$.

B. ϵ -ESs-JamNet-UCB-AILP Algorithm: Unknown CSI at ESs

Due to the online placement of passive jammers, the channel gain $h_{k,j}^t(v)$ is usually unknown and so that the rewards of HES from ESs to the jammer. Now, we consider the practical case that the expected energy rewards are unknown, and online learning algorithms are called for, e.g. UCB [38]. As noticed, it is difficult to combine UCB method directly with the proposed *ALP* for the general heterogeneous-energy-demand systems. However, we can narrow down to a special and very reasonable case, when all jammers have the same energy demand under a given ES, i.e., $c_{k,j} = c_k$ for all j and k , the normalized expected HE $\eta_{k,j}$ represents the quality of jammer j under ES k . In this case, the candidate jammer set for each ES only contains one jammer to be scheduled, which is the jammer with the highest expected HE. Thus, the previous *ALP* algorithm for the known statistics case is simple. When the expected energy rewards are unknown, we can extend the UCB-AILP algorithm by managing the UCB for the normalized expected energy rewards.

When the energy demands for different jammers under the same ES are heterogeneous, it is difficult to combine AILP with the UCB method since the *ALP* algorithm in this case not only requires the ordering of $\eta_{k,j}$'s, but also the ordering of $u_{k,j}$'s and the ratios $\frac{u_{k,j_1} - u_{k,j_2}}{c_{k,j_1} - c_{k,j_2}}$. We propose an ϵ -ESs-JamNet-UCB-AILP Algorithm that explores and exploits separately: the scheduler takes jammers under all ESs in the first $\epsilon(T)$ rounds to estimate the expected energy rewards, and runs *ALP* based on the estimates in the remaining $T - \epsilon(T)$ rounds.

Algorithm 2 ϵ -ESs-JamNet-UCB-AILP

Input: Time horizon T , energy budget \bar{E} , exploration stage length $\epsilon(T)$, and $c_{k,j}$'s, for all k and j ;

```

1: for  $t = 1$  to  $\epsilon(T)$  do
2:   if  $b > 0$  then
3:     Take jammer  $A_t = \operatorname{argmin}_{k \in \mathcal{A}} C_{X_t, j}$  (with random tie-breaking);
4:     Observe the HE  $Y_{A_t, t}$ ;
5:     Update counter  $C_{X_t, A_t} = C_{X_t, A_t} + 1$ ; update remaining energy budget  $b = b - c_{X_t, A_t}$ ;
6:     Update the HE estimate:
       
$$\bar{u}_{X_t, A_t} = \frac{(C_{X_t, A_t} - 1)\bar{u}_{X_t, A_t} + Y_{A_t, t}}{C_{X_t, A_t}}.$$

7:   end if
8: end for
9: for  $t = \epsilon(T) + 1$  to  $T$  do
10:  Remaining time  $\tau = T - t + 1$ ;
11:  if  $b > 0$  then
12:    Obtain the probabilities  $p_{k,j}(b/\tau)$ 's by solving the problem  $(\mathcal{LP}'_{\tau, b})$  with  $u_{k,j}$  replaced by  $\bar{u}_{k,j}$ ;
13:    Take jammer  $j$  with probability  $p_{X_t, j}(b/\tau)$ ;
14:    Remaining energy budget  $b = b - c_{X_t, A_t}$ ;
15:  end if
16: end for

```

For the case of exposition, we assume $c_{k,j_1} \neq c_{k,j_2}$ for any k and $j_1 \neq j_2$ ¹, and let $\Delta_{\min}^{(c)} = \min_{\substack{k \in \mathcal{X} \\ j_1, j_2 \in \{0\} \cup \mathcal{A}}} \{|c_{k,j_1} - c_{k,j_2}|\}$.

Let $\varepsilon_{k,j_1,j_2} = \frac{u_{k,j_1} - u_{k,j_2}}{c_{k,j_1} - c_{k,j_2}}$ for $k \in \mathcal{X}$, $j_1, j_2 \in \{0\} \cup \mathcal{A}$, and $j_1 \neq j_2$ (recall that $u_{k,0} = 0$ and $c_{k,0} = 0$ for the “dummy jammer”), $\bar{\varepsilon}_{k,j_1,j_2}$ be its estimate at the end of the exploration stage, i.e., $\bar{\varepsilon}_{k,j_1,j_2} = \frac{\bar{u}_{k,j_1} - \bar{u}_{k,j_2}}{c_{k,j_1} - c_{k,j_2}}$. Let $\Delta_{\min}^{(\varepsilon)}$ be the minimal difference between any $\varepsilon_{k_1,j_{11},j_{12}}$ and $\varepsilon_{k_2,j_{21},j_{22}}$, i.e.,

$$\Delta_{\min}^{(\varepsilon)} = \min_{\substack{k_1, k_2 \in \mathcal{X} \\ j_{11}, j_{12}, j_{21}, j_{22} \in \{0\} \cup \mathcal{A}}} \{|\varepsilon_{k_1,j_{11},j_{12}} - \varepsilon_{k_2,j_{21},j_{22}}|\}.$$

Moreover, let $\pi_{\min} = \min_{k \in \mathcal{X}} \pi_k$ and let $\Delta^* = \Delta_{\min}^{(c)} \Delta_{\min}^{(\varepsilon)}$. Then, the following lemma states that under ϵ -ESs-JamNet-UCB-AILP with a sufficiently large $\epsilon(T)$, the scheduler will obtain a correct ordering of ε_{k,j_1,j_2} 's with high probability at the end of the exploration stage.

Lemma 5. *Let $0 < \delta < 1$. Under ϵ -ESs-JamNet-UCB-AILP, if*

$\epsilon(T) = \lceil \frac{K}{(1-\delta)\pi_{\min}} + \log T \max\{\frac{1}{\delta^2}, \frac{16K}{(1-\delta)\pi_{\min}(\Delta^)^2}\} \rceil$, then for any contexts $k_1, k_2 \in \mathcal{X}$, and jammers $j_{11}, j_{12}, j_{21}, j_{22} \in \{0\} \cup \mathcal{A}$, if $\varepsilon_{k_1,j_{11},j_{12}} < \varepsilon_{k_2,j_{21},j_{22}}$, then at the end of the $\epsilon(T)$ -th slot, we have*

$$\mathbb{P}\{\bar{\varepsilon}_{k_1,j_{11},j_{12}} < \bar{\varepsilon}_{k_2,j_{21},j_{22}}\} \leq (K+4)T^{-2}.$$

Moreover, the scheduler ranks all the ε_{k,j_1,j_2} 's correctly with probability no less than $1 - (4J+1)KT^{-2}$.

Proof: We first analyze the number of executions for each ES-jammer pair (k, j) in the exploration stage. Let

¹For the case with $c_{k,j_1} = c_{k,j_2}$ for some k and $j_1 \neq j_2$ (and $u_{k,j_1} \neq u_{k,j_2}$), we can correctly remove the suboptimal jammer with high probability by comparing their empirical energy rewards $\bar{u}_{k,j_1} = \bar{u}_{k,j_2}$

$N_k = \sum_{t=1}^{\epsilon(T)} \mathbb{1}(X_t = k)$ be the number of occurrences of ES k up to slot $\epsilon(T)$. Recall that the ESs X_t is activated i.i.d. in each slot by the time-division protocol. Thus, using Hoeffding-Chernoff Bound for each ES k , we have

$$\begin{aligned} & \mathbb{P}\{\forall k \in \mathcal{X}, N_k \geq (1-\delta)\pi_k \epsilon(T)\} \\ & \geq 1 - \sum_{k=1}^K \mathbb{P}\{N_k < (1-\delta)\pi_k \epsilon(T)\} \\ & \geq 1 - Ke^{-2\delta^2 \epsilon(T)} \geq 1 - Ke^{-2\log T} = 1 - KT^{-2}. \end{aligned}$$

On the other hand, the lower bound $(1-\delta)\pi_k \epsilon(T) \geq J + \frac{16J\log T}{(\Delta^*)^2}$, then

$$C_{k,j} \geq \lfloor 1 + \frac{16\log T}{(\Delta^*)^2} \rfloor \geq \frac{16\log T}{(\Delta^*)^2}, \forall j \in \mathcal{A}. \quad (20)$$

Therefore,

$$\mathbb{P}\{\forall k \in \mathcal{X}, \forall j \in \mathcal{A}, C_{k,j} \geq \frac{16\log T}{(\Delta^*)^2}\} \geq 1 - JT^{-2} \quad (21)$$

Next, we study the relationship between the estimates $\bar{\varepsilon}_{k_1,j_{11},j_{12}}$ and $\bar{\varepsilon}_{k_2,j_{21},j_{22}}$ at the end of the exploration stage. We note that

$$\begin{aligned} & \bar{\varepsilon}_{k_1,j_{11},j_{12}} \geq \bar{\varepsilon}_{k_2,j_{21},j_{22}} \\ & \Leftrightarrow (\bar{\varepsilon}_{k_1,j_{11},j_{12}} - \varepsilon_{k_1,j_{11},j_{12}} - \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{2}) \\ & - (\bar{\varepsilon}_{k_2,j_{21},j_{22}} - \varepsilon_{k_2,j_{21},j_{22}} - \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{2}) \geq 0 \\ & \Leftrightarrow (\frac{\bar{u}_{k_1,j_{11}} - u_{k_1,j_{11}}}{c_{k_1,j_{11}} - c_{k_1,j_{12}}} - \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{4}) \\ & - (\frac{\bar{u}_{k_1,j_{12}} - u_{k_1,j_{12}}}{c_{k_1,j_{11}} - c_{k_1,j_{12}}} + \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{4}) \\ & - (\frac{\bar{u}_{k_2,j_{21}} - u_{k_2,j_{21}}}{c_{k_2,j_{21}} - c_{k_2,j_{22}}} + \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{4}) \\ & + (\frac{\bar{u}_{k_2,j_{22}} - u_{k_2,j_{22}}}{c_{k_2,j_{21}} - c_{k_1,j_{22}}} - \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{4}) \geq 0. \end{aligned}$$

Thus, for the event $\bar{\varepsilon}_{k_1,j_{11},j_{12}} \geq \bar{\varepsilon}_{k_2,j_{21},j_{22}}$ to be true, we require that at least one term (with the sign) in the last inequality above is no less than zero. Conditioned on $C_{k,j} \geq \frac{16\log T}{(\Delta^*)^2}$, we can bound the probability of each term according to the Hoeffding-Chernoff bound, e.g., for the first term, we have

$$\begin{aligned} & \mathbb{P}\{\frac{\bar{u}_{k_1,j_{11}} - u_{k_1,j_{11}}}{c_{k_1,j_{11}} - c_{k_1,j_{12}}} - \frac{\varepsilon_{k_2,j_{21},j_{22}} - \varepsilon_{k_1,j_{11},j_{12}}}{4} \geq 0 | C_{k_1,j_{11}} \geq \frac{16\log T}{(\Delta^*)^2}\} \\ & \leq \mathbb{P}\{\bar{u}_{k_1,j_{11}} \geq u_{k_1,j_{11}} + \frac{\Delta^*}{4} | C_{k_1,j_{11}} \geq \frac{16\log T}{(\Delta^*)^2}\} \\ & \leq e^{-2\log T} = T^{-2}. \end{aligned}$$

The conclusion then follows by considering the event $\{C_{k,j} \geq \frac{16\log T}{(\Delta^*)^2}, \forall k \in \mathcal{X}, \forall j \in \mathcal{X}\}$ and its negation. ■

Theorem 6. *Let $0 < \delta < 1$. Under ϵ -ESs-JamNet-UCB-AILP, if*

$$\epsilon(T) \geq \frac{J}{(1-\delta)\pi_{\min}} + \log T \max\{\frac{1}{\delta^2}, \frac{16J}{(1-\delta)\pi_{\min}(\Delta^*)^2}\},$$

then the regret of ϵ -ESs-JamNet-UCB-AILP satisfies:

- if $\rho = \bar{E}/T \neq Q_i$, then $R_{\epsilon\text{-FirstALP}}(T, \bar{E}) = O(\log T)$;
- if $\rho = \bar{E}/T = Q_i$, then $R_{\epsilon\text{-FirstALP}}(T, \bar{E}) = O(\sqrt{T})$.

Proof: (Sketch) The key idea of proving this theorem is considering the event where the ε_{k,j_1,j_2} 's are ranked correctly and its negation. When the ε_{k,j_1,j_2} 's are ranked correctly, we can use the properties of the ALP algorithm with modification on the time horizon and energy budget (subtracting the time and energy budget in the exploration stage, which is $O(\log T)$); otherwise, if the scheduler obtains a wrong ranking results, the regret is bounded as $O(1)$ because the probability is $O(T^{-2})$ and the HE in each slot is bounded. ■

C. A Practical Implementation: Determine $\epsilon(T)$ without Prior Information

In Theorem 6, the scheduler requires the value of Δ^* (in fact $\Delta_{\min}^{(\epsilon)}$ because $\Delta_{\min}^{(c)}$ is known) to calculate $\epsilon(T)$. This is usually impractical since the expected energy rewards are unknown *a priori*. Thus, without the knowledge of $\Delta_{\min}^{(\epsilon)}$, we propose a Confidence Level Estimation (CLE) algorithm for deciding when to end the exploration stage.

Specifically, assume $\Delta_{\min}^{(\epsilon)} > 0$ and is unknown by the scheduler. In each slot of the exploration stage, the scheduler tries to solve the problem $(\mathcal{LP}'_{\tau,b})$ with $u_{k,j}$ replaced by $\bar{u}_{k,j}$ using comparison, i.e., using Algorithm 1 and sorting the virtual jammers. For each comparison, the scheduler tests the confidence level according to Algorithm 3. If all comparisons pass the test, i.e., $\text{flagSucc} = \text{true}$ for all comparisons, then the scheduler ends the exploration stage and starts the exploitation stage.

Algorithm 3 Confidence Level Estimation (CLE)

Input: Time horizon T , estimate $\bar{\epsilon}_{k_1,j_{11},j_{12}}, \bar{\epsilon}_{k_2,j_{21},j_{22}}$, number of executions $C_{k_1,j_{11}}, C_{k_1,j_{12}}, C_{k_2,j_{21}}$, and $C_{k_2,j_{22}}$;
Output: flagSucc ; $\Delta' = \frac{\Delta_{\min(c)}(\bar{\epsilon}_{k_1,j_{11},j_{12}} - \bar{\epsilon}_{k_2,j_{21},j_{22}})}{2}$;
1: **if** $e^{-2(\Delta')^2 \min\{C_{k_1,j_{11}}, C_{k_1,j_{12}}\}} \leq T^{-2} \& e^{-2(\Delta')^2 \min\{C_{k_2,j_{21}}, C_{k_2,j_{22}}\}} \leq T^{-2}$ **then**
2: $\text{flagSucc} = \text{true}$;
3: **end if**
4: **return** flagSucc ;

By similar arguments as in Theorem 4, we can show that the ϵ -First policy with CLE will achieve $O(\log T)$ regret except for the boundary cases, where it achieves $O(\sqrt{T})$ regret. On one hand, according to Hoeffding-Chernoff bound, if all comparisons pass the confidence level test, then with probability at least $1 - KJ^2T^{-2}$, the algorithm obtains the correct rank and provide a right solution for the problem $(\mathcal{LP}'_{\tau,b})$. On the other hand, because $\Delta^* > 0$, from the analysis in the previous section, we know that the exploration stage will end within $O(\log T)$ rounds with high probability. Therefore, the expected regret is the same as that in the case with known $\Delta_{\min}^{(\epsilon)}$.

IV. ONLINE JAMMER PLACEMENT AND POWER ALLOCATION: A SIMPLE ALGORITHM AND A UPPER BOUND

In our online jammer placement model, we receive an unknown number of J friendly jammer placement requests sequentially over time. Each jammer $1 \leq j \leq J$ targets at an eavesdropping location p_e at \mathfrak{F} . Let the short notation $d_{jp_e(j)}$ denote the distance between the jammer j and its jammed eavesdropping location $p_e(j)$. We denote $\Delta = (\max_j d_{jp_e(j)})/(\min_j d_{jp_e(j)})$ as the *distance ratio*. Further, if jammers are informed with a parameter t_j , which denotes the eavesdropping duration within the jamming scope. We $\Gamma = (\max_i t_i)/(\min_i t_i)$ as the *duration ratio*. W.l.o.g., we let $\min_i t_i = 1$ and $\max_i t_i = \Gamma$. Jammers arrive sequentially over time and the goal is to accept the minimal number of

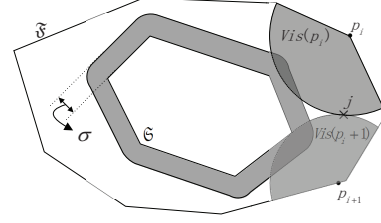


Fig. 3: Guarded safe distance σ for online jammer placement.

requests to interfering all potential eavesdropper while making minimal noises to legitimate communications.

For each jammer placement request, an online algorithm must decide whether to accept the request or deny it, and the decision can not be revoked. For an accepted jammer j it needs to set a power level P_j and a channel $f_j \in \{1, \dots, F\}$ to it to emit the interfering power. In the following we first analyze the spatial aspect of the problem and assume that eavesdropping behavior last forever, i.e., $t_j = \infty$. We begin by analyzing a simple online algorithm for the case of a single channel and any polynomial power assignment. Our analysis of the online algorithm introduces a number of critical observations that are used in later subsections.

The main idea of the algorithm is to accept a new jammer only if it keeps a *safe distance* σ from every other previously accepted jammers to meet two goals: 1) the sum of cumulated interfering power of all jammers to any legitimate communications are small enough; 2) the maximized safe distance σ make the number of jammers placed to be minimized. In particular, we accept incoming jammer i only if $\min\{d_{j,p_s}, \forall p_s \in \partial\mathfrak{S}, \forall j \in J\} \geq \sigma$ and $\max\{d_{i,p_e(j)}, d_{j,p_e(i)}\} \geq \sigma$ for every other previously accepted jammer $j \in J$. We call this algorithm JAM-SAFE-DISTANCE. There is a important tradeoff for the choice of σ among EE of interfering power, validity and competitive ratio. A larger σ means safeguarding a large scope of eavesdropping locations with high interfering power and have resulted in minimized number of jammers to place, but the SIR constraints of legitimate communications might become violated. If σ is too small, then more jammers are required to be placed and at some point the accumulated interference at an accepted jammer placement request can get too large and the SIR constraint of legitimate communications becomes violated.

We strive to devise the JAM-SAFE-DISTANCE to make σ as large as possible to ensure optimal competitive ratio as well as not to violate SIR constraints. On the one hand, we need to bound the interference on the edge of the fence \mathfrak{S} at accepted jammer placement requests to construct a worst-case legitimate communication scenario. On the other hand, we consider an accepted jammer j block an eavesdropper $p_e(j)$ with certain distance r_j . In the following we show that for $r \in [0, 1]$ the choice of

$$\sigma = \min \left\{ 2\Delta, \max \left\{ 4\Delta^r \sqrt{\frac{72\delta_s}{P(\gamma-2)}}, \Delta^{(1-r)} \sqrt{\frac{\tilde{P}}{\delta_e}} \right\} \right\} \quad (22)$$

is sufficient to yield the Theorem 5 in the following. Denote the $\underline{L} = \min \|\mathfrak{S} - \mathfrak{F}\|$ as the minimal distance among \mathfrak{S} and \mathfrak{F} . If the eavesdropping locations on the \mathfrak{F} are unavailable, we

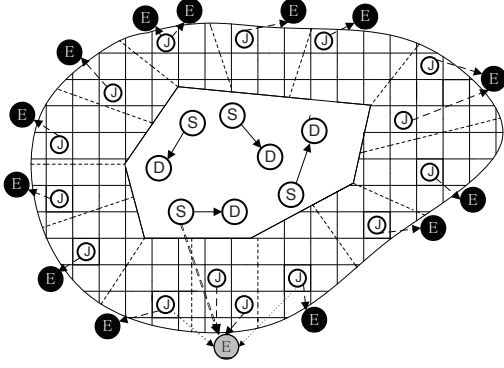


Fig. 4: Interference contributed from each segmented Sectors.

need the additional condition that

$$\underline{L} \geq (\sqrt{2} + 2)\sigma \quad (23)$$

to block all eavesdropping points on \mathfrak{F} .

Theorem 7. *In a single channel scenario, JAM-SAFE-DISTANCE is $\Omega(\Delta)$ -competitive for any polynomial interfering power assignment with $r \in [0, 1]$.*

Proof: We first show that JAM-SAFE-DISTANCE is valid, i.e., for an accepted jammer j the SIR constraint of any locations at the edge of \mathfrak{S} never becomes violated. In particular, we will underestimate the distances of already accepted jammers to overestimate the interference at any position $p_s, \forall p_s \in \partial\mathfrak{S}$. As such, even under the worst conditions the SIR constraint at any potential legitimate receiver, i.e., p_s , will remain valid.

To estimate the interference at p_s , we have to calculate how many jammers may be placed at which distance. Using the fact $\underline{L} \geq 2\Delta \geq \sigma$ as shown in Fig. 3. Then, it is straightforward to devise the rule of the algorithm such that any two different accepted jammers are at least a distance of $\sigma - \Delta \leq \sigma/2$ apart to block eavesdroppers on the \mathfrak{F} . We segment all of \mathbb{R}^2 into 2-dimensional squares with length $\sigma/4$ and we call it *sectors*. The greatest distance within a sector is $\sigma\sqrt{2}/4 = \sigma/2\sqrt{2} \leq \sigma/2$. Each sector can contain jammer from at most one request, so there are at most two jammers in every sector.

W.l.o.g., we assume that sectors are created such that the jammer j lies in a corner point of 2^2 sectors. We divide the set of sectors into *layers*. The first layer consists of the 2^2 sectors incident to j . The second layer are all sectors not within the first layer but share at least a point with sector from the first layer, and so on. Hence, there are $(2l)^2$ sectors from layers 1 through l , and their union is a large square of side length $2l\sigma/4$ with j in the center. Therefore, there are exactly $2^2(l^2 - (l-1)^2)$ sectors in layer l . Due to the algorithm there can be no sender at a distance smaller than σ from j . The sector of smallest layer that is at a distance at least σ from j can be reached along the diagonal of the squares of that layer. There can be no jammer in all sectors from layers 1 through l' , where l' is bounded by $\sigma \leq l'(\sigma/2\sqrt{2})$, which yields $l' \geq 3$. For bounding the interference assume that in all sectors of layer $l \geq 3$ there are two jammers. Note that all jammer in sectors from a layer l have a distance at least $(l-1)\sigma/4$ to j . To bound the interference that is created at j , we use the following technical lemma from [36] under \mathbb{R}^2 . ■

Lemma 8. *For $\gamma > 2 \geq 1$ the following holds:*

$$2^2 \cdot \sum_{l=3}^{\infty} \frac{l^2 - (l-1)^2}{(l-1)^\gamma} < \frac{36}{\gamma-2}.$$

With Lemma 6 and set $P_j = d_{jp_e(j)}^{r\gamma}$, we bound the interference for legitimate communication

$$I = \sum_{j \in J} \frac{d_{jp_e(j)}^{r\gamma}}{d_{jp_s}^\gamma} < 2\Delta^{r\gamma} \sum_{l=3}^{\infty} \frac{2^2(l^2 - (l-1)^2)}{((l-1)\sigma/4)^\sigma} < 2\Delta^{r\gamma} \left(\frac{4}{\sigma}\right)^\gamma \cdot \frac{36}{\gamma-2}.$$

To satisfy the SIR constraint at p_s , we let $\bar{P} \geq \delta_s I$, i.e.,

$$2\delta_s \Delta^{r\gamma} \cdot \left(\frac{4}{\sigma}\right)^\gamma \cdot \frac{36}{\gamma-2} \leq \bar{P}.$$

This yields a lower bound for the distance σ ,

$$\sigma \geq 4 \cdot \Delta^r \cdot \sqrt[\gamma]{\frac{72\delta_s}{P(\gamma-2)}}, \quad (24)$$

which can be verified to hold for our choice of σ .

Then, to bound the SIR constraint at p_e , we use the fact that in the worst condition if there is no other jammer, a single jammer j is enough to thwart the eavesdropper, i.e., $d_{jp_e(j)}^{r\gamma}/d_{jp_e(j)}^\gamma \delta_e \geq \bar{P}/d_{s(p_e)p_e}^\gamma$. Note that $d_{jp_e(j)}^{r\gamma}/d_{jp_e(j)}^\gamma = d_{jp_e(j)}^{(r-1)\gamma} \geq \Delta^{(r-1)\gamma}$ and $d_{s(p_e)p_e} \geq \underline{L} \geq \sigma$, we have

$$\Delta^{(r-1)\gamma} \delta_e \geq \frac{\bar{P}}{\sigma^\gamma}. \quad (25)$$

This yields another lower bound for σ

$$\sigma \geq \Delta^{(1-r)} \sqrt[\gamma]{\frac{\bar{P}}{\delta_e}}. \quad (26)$$

Combine results (24) and (27) yields (22).

Moreover, when the locations of eavesdroppers are unavailable, every placed jammer is necessary to block all eavesdropping position on the intersections of its sector and the fence \mathfrak{S} . Denote the furthest eavesdropping position within the sector as p'_e . Then we have $d_{jp_e(j)}^{r\gamma}/d_{jp_e(j)}^\gamma \delta_e \geq d_{jp_e(j)}^{r\gamma}/(d_{jp_e} + d_{p'_e p_e})^\gamma \delta_e$. Due to the size of the sector we have that $d_{p'_e p_e} \leq \sqrt{2}$. Also $d_{jp_e} \geq 1$, which implies

$$\frac{d_{jp_e(j)}^{r\gamma} \delta_e}{(d_{jp_e} + d_{p'_e p_e})^\gamma} \geq \frac{1}{(\sqrt{2}+1)^\gamma} \frac{d_{jp_e(j)}^{r\gamma} \delta_e}{d_{jp_e}^\gamma} \geq \frac{\Delta^{(r-1)\gamma} \delta_e}{(\sqrt{2}+1)^\gamma}. \quad (27)$$

On the other hand, use the condition (23) that we upper bound $\bar{P}/d_{s(p'_e)p'_e}^\gamma \leq \bar{P}/(\underline{L}-\sigma)^\gamma \leq \frac{\bar{P}}{(\sqrt{2}+1)^\gamma \sigma^\gamma}$. Use the fact that (25), we have

$$\frac{\Delta^{(r-1)\gamma} \delta_e}{(\sqrt{2}+1)^\gamma} \geq \frac{\bar{P}}{(\sqrt{2}+1)^\gamma \sigma^\gamma}. \quad (28)$$

To bound the competitive ratio we need the following *Density Lemma*, which is motivated by Lemma 3 in Andrews and Dinitz [35] to restrict interference both from senders and at receivers for any legitimate communication links. However, the placement of friendly jammer is interesting to be found as a different problem. In this case, we need to estimate the interference caused at the legitimate communication of a placed jammer by mapping its transmission power calculated from the SIR constraint at the eavesdropper.

Lemma 9. (*Density Lemma*) *Assume a sector A with side-length $x \geq 1$ and any feasible jammer placement solution with arbitrary power assignment. There can be only $\frac{\delta_e 3^\gamma \bar{L}^\lambda}{\delta_s} \frac{\bar{P}}{P} (x+1)^2$ jammer placement requests in A .*

Proof: We first assume $x = 1$ and consider the number of jammers in section A . At first, the interfering power receiving

have length Δ . From the SIR constraints, we bound the minimum distance every other successful jammer request has to keep to the fence \mathcal{F} . This yields a blocked area in which the online jammer placement algorithm is not able to accept any request. We then count the maximum number of requests that can be placed into this region, where the optimum solution can accept simultaneously.

To extend the previous arguments to arbitrary distance-based power assignments, we observe that the previous lower bound uses only requests of length 1 and Δ . Let ϕ be the function of the distance-based power assignment, then $\phi(\Delta)$ is the power of the first request. The lower bound for this power assignment behaves exactly as for a polynomial assignment with $r = (\log \phi(\Delta)) / (\alpha \log \Delta)$.

Note that when a power assignment is not distance-based, it might assign different powers to small requests based on whether they are near the sender or the receiver of the first request. This is not helpful since the jammer have a direct interfering link to the eavesdroppers. In this case, we create the same instance using only undirected requests. Then we get a blocked area of at least $\Omega(\Delta)$ for any polynomial power assignment around both points of the first request. Using the normalization of powers as before we observe that there is a blocked area of size $\Omega(\Delta)$ for any small request, *no matter which power we assign to it*. This proves the theorem. ■

V. IMPROVED COMPETITIVE RATIOS AND EE UNDER SPATIAL AND TEMPORAL EXTENSIONS

A. Jammer placement request with duration

In the previous sections we assumed that requests last forever, analyzing only the spatial aspect of the problem. We now show how our results extend when each request i has a duration t_i . After time t_i an accepted request stops sending and leaves (thus, no longer causing interference).

We show the modification for the algorithm SAFE-DISTANCE for $r \in [0, 1]$. We adapt the algorithm in the following way. It accepts a given request i if and only if the safe distance σ holds to all previously accepted requests that are active at some point time in i 's duration.

Our first observation is the following. If we consider a fixed point in time, an optimal solution OPT can have at most $O(\Delta^d)$ more requests than our algorithm, as this corresponds to the spatial problem. Now let i be a request accepted by SAFE-DISTANCE with smallest duration possible, that is, $t_i = 1$. Each request contained in an optimal solution that interferes with i is active at least either when i starts or when it stops sending. So it is sufficient to count the accepted requests in OPT at both of these points in time to upper bound the number of requests blocked by i , which is $2 \cdot O(\Delta^d)$. Furthermore, a request i with $t_i \leq \Gamma$ can be split into at most Γ requests of duration 1, thus blocking at most $(\Gamma + 1) \cdot O(\Delta^d)$ requests. The argumentation is similar for other polynomial power assignments and results in an additional factor of Γ in all previously shown bounds.

In the case of multiple channels, for $k = k' \cdot k''$, clustering of requests w.r.t. similar length and duration values can be used to improve the ratio for our algorithm SAFE-DISTANCE to $O(k \cdot$

$\Gamma^{1/k'} \Delta^{(d/2k'')+\epsilon})$. Choosing $k = \log \Gamma \cdot \log \Delta$, RANDOMSAFE-DISTANCE becomes $O(\log \Gamma \cdot \log \Delta)$ -competitive.

B. Multiple Channels

In this section we show how to generalize the algorithms above to k channels and decrease their competitive ratio. We propose a k -channel adjustment, in which we separate the problem by using certain channels only for specific request lengths. All requests with length in $[\Delta^{(i-1)/k}, \Delta^{i/k}]$ are assigned to channel i , for $i = 1, \dots, k$, where we assign requests of length $\Delta^{i/k}$ arbitrarily to channel i or $i+1$. For each channel i we apply an algorithm outlined above, which makes decisions about acceptance and power of requests assigned to channel i . Using this separation, we effectively reduce the aspect ratio to $\Delta^{1/k}$ on each channel. If the optimum solution has to adhere to the same length separation on the channels, this would yield a denominator k in the exponents of Δ of the competitive ratios. Obviously, the optimum solution is not tied to our separation, but the possible improvement due to this degree of freedom can easily be bounded by a factor k . This yields the following corollary.

Corollary 12. *MULTI-CLASS SAFE-DISTANCE with k -channel adjustment is $O(k \Delta^{(d/2k)+\epsilon})$ -competitive using the square-root power assignment. SAFE-DISTANCE with k -channel adjustment is $O(k \Delta^{d/k})$ -competitive for any polynomial assignment with $r \in [0, 1]$, and $O(k \Delta^{\max\{r, 1-r\} \cdot d/k})$ -competitive for $r \notin [0, 1]$.*

VI. SIMULATIONS

We conducted preliminary experiments to compare the different numbers of the eavesdroppers. The setting we have chosen is the storage/fence shown in Figure 5. The fence is of dimensions 500×300 units and we placed a grid of 1×1 cells in the entire region. We simulated both JAM-SAFEDIST-POWER and JAM-LIFEMAX in this setting. For the power assignment from JAM-SAFEDIST-POWER, we investigated the difference in number of jammers. Finally, we observed the variation in total power assigned with ϵ and δ and the number of jammers placed with ϵ , δ and \hat{P} . We set the round number T is 300. Give the energy conversion rate $\alpha(v) = 0.4$ and the average distance $\bar{l} = 4$ as the distance $l \in [1, 10]$. Therefore we can conclude the energy attenuation factor $\bar{l}^\gamma = 16$ when the $\gamma = 4$. We chosen the following values: (i) $\epsilon = \{0.1, 0.2, 0.3, 0.4, 0.5\}$, (ii) $\delta = \{0.5, 0.6, \dots, 1.0\}$, (iii) $\hat{P} = \{(1/\epsilon), (2/\epsilon), \dots, (5/\epsilon)\}$. In both numbers of the eavesdroppers, we removed all grid points which were in the forbidden region.

For JAM-SAFEDIST-POWER, the decline is more steep than the JAM-LIFEMAX because of the 300 rounds cause the JAM-LIFEMAX data is obtained by repeated average. The JAM-SAFE-DistPower and JAM-LIFE-Max give the same information about the desired jammers become more and more as the eavesdroppers goes up. The last two figures about the JAM-SAFEDIST-MultiChannel talk us the multichannel give the much less desired jammers but the benefits of multichannel will reduce as the channel grows in number in two different ways based on the $N_e = 6$.

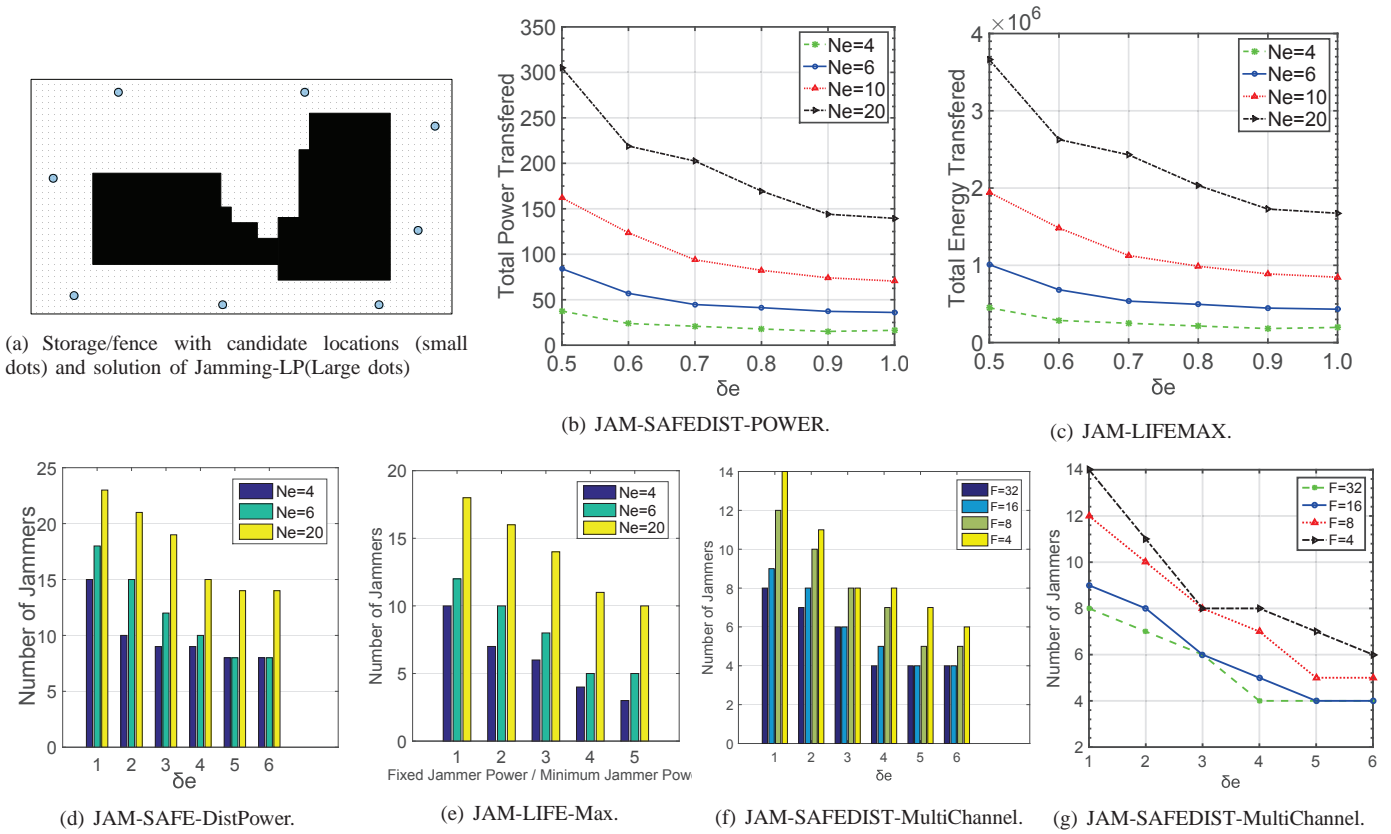


Fig. 6: Results of simulations under proposed online algorithms.

VII. CONCLUSION

In this paper, we propose the first distributed protocol that provides secure communication in any geographically restricted communication networks using energy-constrained friendly jammers wirelessly powered by legitimate transmitters as energy sources. Online learning algorithms are proposed to maximize the lifetime of jammer and met the goal of EE with heterogenous energy demands. Our protocol supports dynamic behaviors, e.g., mobility, eavesdropping (communicating) completion or addition/removal of nodes, as along the secure communication are restricted to the storage. However, our proposed protocol is adaptive to the situations such information is available, e.g., exact positions and frequency of both legitimate communications and eavesdropping behaviors, and foreseen further EE improvements and reduced number of jammers. We provided competitive ratios for approximate algorithms in several distributed settings, and found the multi-channel diversity is a good approach to improve the security of wireless communications.

REFERENCES

- [1] A. Mukherjee, S. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550-1573, Aug. 2014.
- [2] X. Zhou, L. Song, and Z. Zhang, *Physical Layer Security in Wireless Communications*, CRC Press, 2013.
- [3] X. Tang, R. Liu, P. Spasojevic, and H. Poor, "On the throughput of secure hybrid-ARQ protocols for gaussian block-fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1575-1591, Apr. 2009.
- [4] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180-2189, Jun., 2008.
- [5] J. Ko, C. Lu, M. B. Srivastava, J. Stankovic, A. Terzis, and M. Welsh, "Wireless sensor networks for healthcare. Proc. of the IEEE," vol. 98, no. 11, pp. 1947-1960, 2010.
- [6] M. Loh, and A. Tam, "Wireless smart card and integrated personal area network, near field communication and contactless payment system. U.S. Patent Application 12/234,499, 2008.
- [7] Y. Xiao, X. Shen, B. O. Sun, and L. Cai, "Security and privacy in RFID and applications in telemedicine," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 64-72, 2006.
- [8] F. Hu, and N. K. Sharma, "Security considerations in ad hoc sensor networks. *Ad Hoc Networks*," vol. 3, no. 1, pp. 69-89, 2005.
- [9] E. Arkin, Y. Cassuto, A. Efrat, G. Grebla, J. S. Mitchell, S. Sankararaman, and M. Segal, "Optimal placement of protective jammers for securing wireless transmissions in a geographic domain," In *Proc. of the 14th International Conference on Information Processing in Sensor Networks (IPSN' 15)*, pp. 37-46, 2015.
- [10] S. Sankararaman, A. K. Abu-Affash, A. Efrat, S. D. Eriksson-Bique, V. Polishchuk, S. Ramasubramanian, and M. Segal, "Optimization schemes for protective jamming," In *13th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc'12*, pp. 65-74, 2012.
- [11] Y. Allouche, Y. Cassuto, A. Efrat, M. Segal, E. Arkin, G. Grebla, and J. S. Mitchell, "Secure Communication through Jammers Jointly Optimized in Geography and Time," In *16th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc'15*, pp. 112-123, 2015.
- [12] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun., pp. 2363-2367, 2010.
- [13] L. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1612-1616.
- [14] L. Liu, R. Zhang, and K. C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 288-300, 2013.
- [15] S.Z. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication:

- opportunities and challenges,” IEEE Communications Magazine, vol.53, no.4, pp. 117-125, 2015.
- [16] L. Liu, R. Zhang, and K.-C. Chua, “Wireless information transfer with opportunistic energy harvesting,” IEEE Trans. Wireless Commun., vol. 12, no. 1, pp. 288-300, Jan. 2013.
 - [17] R. Zhang and C. K. Ho, “MIMO broadcasting for simultaneous wireless information and power transfer,” IEEE Trans. Wireless Commun., vol. 12, no. 5, pp. 1989-2001, May 2013.
 - [18] X. Zhou, R. Zhang, and C. K. Ho, “Wireless information and power transfer: Architecture design and rate-energy tradeoff,” IEEE Trans. Commun., vol. 61, no. 11, pp. 4754-4767, Nov. 2013.
 - [19] H. Ju and R. Zhang, “Throughput maximization in wireless powered communication networks,” IEEE Trans. Wireless Commun., vol. 13, no. 1, pp. 418-428, Jan. 2014.
 - [20] H. Xing, L. Liu, and R. Zhang, “Secrecy wireless information and power transfer in fading wiretap channel,” in Proc. IEEE Int. Conf. Commun., pp. 5402-5407, Jun. 2014.
 - [21] L. Liu, R. Zhang, and K.C. Chua, “Secrecy wireless information and power transfer with MISO beamforming,” IEEE Trans. Signal Process., vol. 62, no. 7, pp. 1850-1863, Apr. 2014.
 - [22] D. Ng, E. Lo, and R. Schober, “Robust beamforming for secure communication in systems with wireless information and power transfer,” IEEE Trans. Wireless Commun., vol. 13, no. 8, pp. 4599-4615, Aug. 2014.
 - [23] R. Feng, Q. Li, Q. Zhang, and J. Qin, “Robust secure transmission in MISO simultaneous wireless information and power transfer system,” IEEE Trans. Veh. Technol., vol. 64, no. 1, pp. 400-405, May, 2014.
 - [24] W. Liu, X. Zhou, S. Durrani, and P. Popovski, “Secure Communication with a Wireless-Powered Friendly Jammer”, IEEE Trans. Wireless Commun., vol. 15, no. 1, pp. 401-415, 2016.
 - [25] A. Mukherjee, A. L. Swindlehurst, “Detecting passive eavesdroppers in the MIMO wiretap channel,” In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2809-2812, 2012.
 - [26] X. Zhou, R. Zhang, C. K. Ho, “Wireless information and power transfer: Architecture design and rate-energy tradeoff,” IEEE Transactions on Communications, vol.61, no.11, pp. 4754-4767, 2013.
 - [27] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, “Wireless-powered relays in cooperative communications: Time-switching relaying protocols and throughput analysis,” IEEE Trans. Commun., vol. 63, no. 5, pp. 1607-1622, 2015.
 - [28] Wyner, A. D. The wire-tap channel. Bell System Technical Journal, The, vol.54, no.8, pp. 1355-1387, 1975.
 - [29] X. Zhang, X. Zhou, and M. McKay, “On the design of artificial-noiseaided secure multi-antenna transmission in slow fading channels,” IEEE Trans. Veh. Technol., vol. 62, no. 5, pp. 2170-2181, Jun. 2013.
 - [30] S. A. Fakoorian and A. L. Swindlehurst, “Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer,” IEEE Transactions on Signal Processing, vol.59, no.10, pp.5013-5022, 2011.
 - [31] A. Mukherjee and J. Huang, “Deploying multi-antenna energy-harvesting cooperative jammers in the MIMO wiretap channel,” in Proc. IEEE Conf. Rec. 46 Asilomar Signals Syst. Comput., pp. 1886-1890, 2012.
 - [32] J. P. Vilela and J. Barros, “Collision-free jamming for enhanced wireless secrecy,” In IEEE 14th International Symposium on A World of Wireless, Mobile and Multimedia Networks, WoWMoM 2013, pp.1-6, 2013.
 - [33] J. P. Vilela, P. C. Pinto, and J. Barros, “Jammer selection policies for secure wireless networks,” In IEEE International Conference on Communications (ICC) Workshops, pp. 1-6, 2011.
 - [34] S. Bubeck and N. Cesa-Bianchi, “Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems,” Foundation and Trends in Machine Learning, vol. 5, 2012.
 - [35] M. Andrews, and M. Dinitz, “Maximizing capacity in arbitrary wireless networks in the SINR model: Complexity and game theory,” In Proc. 28th IEEE conf. computer communications (INFOCOM), 2009.
 - [36] Fanghänel, A., Geulen, S., Hofer, M., and B. Vöcking, “Online capacity maximization in wireless networks,” Journal of scheduling, vol.16, no.1, 81-91, 2013.
 - [37] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.
 - [38] P. Auer, N. Cesa-Bianchi, and P. Fischer, “Finite-time analysis of the multiarmed bandit problem,” Machine learning, vol. 47, no. 2, pp. 235-256, 2002.

